Review of Software Packages for Bayesian Multilevel Modeling

Yujiao Mai

University of Central Florida

Zhiyong Zhang

University of Notre Dame

Draft of 01/07/18

Citation: Mai, Y., & Zhang, Z. (2018). Review of software packages for Bayesian multilevel modeling. Structural Equation Modeling, 25(4), 650–658. https://doi.org/10.1080/10705511.2018.1431545

Author Note

This research was supported by a grant from the U.S. Department of Education (R305D140037). There is no conflict of interest.

Correspondence concerning this article should be addressed to Yujiao Mai at maiyujiao@gmail.com or Zhiyong Zhang at zhiyongzhang@nd.edu, Department of Psychology, Notre Dame, IN 46556, USA.

Abstract

Multilevel modeling is a statistical approach to analyze hierarchical data, which consist of individual observations nested within clusters. Bayesian methods is a well-known, sometimes better, alternative of Maximum likelihood methods for fitting multilevel models. Lack of user-friendly and computationally efficient software packages or programs was a main obstacle in applying Bayesian multilevel modeling. In recent years, the development of software packages for multilevel modeling with improved Bayesian algorithms and faster speed has been growing. This article aims to update the knowledge of available software packages for Bayesian multilevel modeling and therefore to promote the use of these packages. Three categories of software packages capable of Bayesian multilevel modeling including brms, MCMCglmm, glmmBUGS, Bambi, R2BayesX, BayesReg, R2MLwiN and others are introduced and compared in terms of computational efficiency, modeling capability and flexibility, as well as user-friendliness. Recommendations to practical users and suggestions for future development are also discussed.

Keywords: Bayesian computer software, MCMC, multilevel modeling, R packages

Review of Software Packages for Bayesian Multilevel Modeling

Multilevel modeling (also known as analysis of hierarchical or mixed-effects models) is a statistical approach to analyze hierarchical data that consist of individual observations nested within clusters/groups/sites (Goldstein, 2011; Kreft, Kreft, & de Leeuw, 1998). Common examples of hierarchical data include clustered cross-sectional data (Broström & Holmberg, 2011), longitudinal data (Wang, 2016), repeated-measures data (Goldstein, Healy, & Rasbash, 1994), and spatial data (Banerjee, Carlin, & Gelfand, 2014). The most well-known techniques for fitting multilevel models include variations of maximum likelihood (ML) and empirical Bayesian estimators (Goldstein, 2011). ML-based methods, with sophisticated implementation in software packages (see Bates, Mächler, Bolker, & Walker, 2014; Peugh & Enders, 2005; West & Galecki, 2011), have been frequently employed in applications (Dedrick et al., 2009). While Bayesian methods are still less widely used (Dedrick et al., 2009), it offers a number of advantages, such as small sample size requirement (Austin, 2010; Dedrick et al., 2009), flexibility to specify complex models like non-normal random-effect models (Lee & Thompson, 2008; Zhang, 2016), and benefits from using empirical information (e.g., Harbord, Whiting, et al., 2009; Zhang, Hamagami, Wang, Nesselroade, & Grimm, 2007). Major obstacles in applying Bayesian multilevel modeling include its time-consuming process and lack of user-friendly software packages or programs (Dedrick et al., 2009; Karabatsos, 2017; Rue et al., 2017).

In recent years, the steady improvement of computer hardware speed and efficiency of Bayesian algorithms (e.g., integrated nested Laplace approximation; Rue, Martino, & Chopin, 2009) as well as the attentiveness to Bayesian estimation of complex multilevel models (e.g., Aguero-Valverde, 2013; Chagneau, Mortier, Picard, & Bacro, 2011; Han & Chaloner, 2004; Pang, 2010) have inspired a surge in the development of software packages for Bayesian

multilevel modeling (e.g., Buerkner, 2017; Correia, Tura, & Lanzillotti, 2016; Grant, Carpenter, Furr, Gelman, et al., 2017; Leckie, Charlton, et al.,2013; Rue et al., 2017). In this article, we aim to review the available software packages for Bayesian multilevel modeling and therefore hope to promote the use of these packages in scientific research. In the following, we first introduce three categories of software packages and compares their technical features and user-friendliness. Then, we provide recommendations to practical users for the choice of difference packages and discuss directions for future development.

Three Categories of Software Packages

According to the analysis purposes we group the software packages available for Bayesian multilevel modeling into three categories: Category A for general purpose Bayesian analysis, categorical B for general purpose Bayesian multilevel modeling, and category C for a particular type of multilevel models.

Category A: General Purpose Software Packages

The software packages in category A are general purpose ones for Bayesian analysis not limiting to certain types of models. The BUGS projects (Gilks, Thomas, & Spiegelhalter, 1994) including WinBUGS (D. Spiegelhalter, Thomas, Best, & Lunn, 2003) and OpenBUGS (D. Spiegelhalter, Thomas, Best, & Lunn, 2014) along with the corresponding R interface R2WinBUGS/R2OpenBUGS (Sturtz, Ligges, & Gelman, 2005) and a web interface WebBUGS (Zhang, 2014) have been the most well-known and widely used in the last few decades. Similar software programs include JAGS (Plummer, 2017) with the R interface rjags (Plummer, Stukalov, & Denwood, 2016), SAS MCMC (Chen, Brown, & Stokes, 2016), Stan (Luo & Jiao, 2017) along with the R interface rstan (Guo et al., 2017), and R–INLA (or INLA; Lindgren & Rue, 2015; Rue, Martino, Lindgren, Simpson, & Riebler, 2013). LaplacesDemon (Hall, 2016)

and MultiBUGS (Goudie, Turner, De Angelis, & Thomas, 2017) are the most recent published packages in this category.

Popularity. Table 1 presents the utilization of various software packages in Bayesian multilevel modeling in the recent years. The results were based on a collection of 92 applied studies between 2012 and 2016. About 70% of the studies used general purpose Bayesian software packages. About 26% of the studies developed their own programs to perform the analyses. BUGS and self-implemented R programs were the dominant tools employed in the reviewed studies. The data in the table also suggested that Matlab programs were the major tool used in neuroscience studies.

Computational efficiency. Each alternative of BUGS was developed with aims to improve the Bayesian algorithms and the computing speed, either by optimizing the algorithms themselves or implementing the algorithms in a faster language (see Table 2 for details). Various Markov Chain Monte Carlo (MCMC) algorithms, such as the Gibbs sampler (Gelfand, 2000), Metropolis–Hastings (Chib & Greenberg, 1995), and slice samplings (Neal, 2003) are built in BUGS. JAGS implements the BUGS algorithms with C++ language to improve the computing speed. Stan (rstan) implements two computationally more efficient MCMC algorithms, Hamiltonian Monte Carlo (HMC; Neal, 1993) and No-U-Turn sampler (NUTS; Hoffman & Gelman, 2014). Due to the nature of MCMC algorithms, however, as the model complexity and the sample size increase, the computing time increases dramatically. Thus, to reduce the running time of MCMC, MultiBUGS implements an algorithm to allow parallel computing for sampling a single chain (Goudie et al., 2017), which is a good solution for researchers having access to high-performance computers or systems. In a similar way, R-INLA adopts a non-MCMC algorithm to approximate Bayesian inference, the integrated nested Laplace approximation

(INLA) approach. In general, the INLA approach can provide good or even exact approximation while reducing computational costs substantially (Rue et al., 2017).

Supported models and priors. In general, the general-purpose Bayesian software packages allow users to specify "unlimited" types of multilevel models and to flexibly customize the parameter priors of various distributions (Lunn, Spiegelhalter, Thomas, & Best, 2009). Nevertheless, this is not true for packages using Laplace approximation such as R-INLA. Although R-INLA supports a large number of models (e.g., latent models including several spatial models) and allow specifying more complex models (Gómez-Rubio & Rue, 2017), R-INLA is not as flexible as BUGS or other MCMC software packages in using complex hierarchical prior structures or handling models with a large number of hyper-parameters (Carroll et al., 2015; Umlauf, Adler, Kneib, Lang, & Zeileis, 2015).

User-friendliness. A fundamental reason behind the popularity of general purpose Bayesian software packages is their flexibility to estimate many different kinds of models. A disadvantage coming with the flexibility is the steep learning curve for users who do not have a good understanding of Bayesian methodology. First, many of these software packages do not provide default model templates or prior types (as shown in Table 3). Therefore, users need to specify the models in detail. Since any model can be specified regardless of whether it makes sense or not, when the models are not specified as expected the users may not realize the mistakes. Second, many of these software packages do not offer optimized algorithms for different models, users are required to understand the features of different algorithms and make a choice. Among the software packages in category A, R-INLA is an exception in that it provides routine functions as well as default prior types and values. Thus, it is relatively user-friendly. Importantly, all the software packages in this category have outstanding documentations and

good maintenance to aid the learning for new users.

Category B: General Purpose Bayesian Multilevel Modeling Software Packages

Software packages in Category B can be used to analyze a variety of multilevel models using Bayesian methods. They include brm with alternative versions rstanarm and rethinking (see comparisons by Buerkner, 2017), MCMCglmm (Hadfield et al., 2010), glmmBUGS (Brown & Zhou, 2010), glmmAK (Komárek & Lesaffre, 2008), blme (Dorie, 2015), and Python Bambi (Yarkoni & Westfall, 2016). Other packages consist of functions or options for using Bayesian methods to fit multilevel models: R2BayesX (BayesX; Umlauf et al., 2015), MCMCpack (Martin, Quinn, & Park, 2011), DPpackage (Jara, Hanson, Quintana, Mueller, & Rosner, 2017), Matlab BayesReg (Karabatsos, 2017), Stata bayesmh (Grant, Furr, Carpenter, & Gelman, 2016), bayesm (Rossi, 2017), MLwiN (Browne, 2017) with R interface R2MLwiN (Zhang, Parker, Charlton, & Browne, 2016), glmmADMB (Bolker, Skaug, Magnusson, & Nielsen, 2016), Mplus (Asparouhov & Muthén, 2010), and arm (Gelman et al., 2016).

Computational efficiency, supported models, priors and outputs. The computational efficiency of these software packages is inherited from their Bayesian engines. Whether they support parallel computing is also a factor influencing their computational efficiency (see Table 2). For example, brms that is based on Stan is more computationally efficient than MCMCglmm (available in Stata) as shown in a simulation study (Buerkner, 2017). Another simulation study (Grant et al., 2016) showed that the Stan Bayesian engine also outperforms Stata’s bayeshm.

In terms of model specification, these software packages are generally not as flexible as those in category A. Table 4 lists the major types of multilevel models supported by each package. Most of these software packages only support a limited number of models. Among these software packages, brms, MCMCglmm, R2BayesX, R2MLwiN, and Mplus Bayes support

almost all the common types of multilevel models. With additional capability, R2BayesX can handle generalized additive mixed models with a large number of parameters and large datasets (e.g., more than 1000 parameters and 200,000 observations; see Brezger & Lang, 2006). While packages, such as glmmBUGS, Bambi, MCMCpack, bayesmh and R2MLwiN allow users to customize a new model through writing scripts. For prior options, brms, Bambi, and R2MLwiN support various types of distributions, while others only support one type or a few types of distributions.

Most of the software packages can output regular summary statistics and plots for convergence diagnostics. For goodness of fit indices, MCMCglmm, R2BayesX, MCMCpack, R2MLwiN, and arm output the deviance information criterion (DIC; D. J. Spiegelhalter, Best, Carlin, & Van Der Linde, 2002), while brms and Bambi both provide Watanabe-Akaike information criterion (WAIC; Watanabe, 2010) as well as Bayes' factors (Kass & Raftery, 1995).

User-friendliness. Different from the software packages in category A, the software packages in category B are relatively easy to use for applied researchers (see Table 3). First, all these software packages provide default model templates or routine functions for the convenience of users. Some of these software packages employ lme4 or lme4-like formulas to specify models, which is definitely an advantage for users with experience using lme4, the most widely used R package for multilevel modeling. Particularly, Matlab's BayesReg has a graphical interface that allows users to specify a model through interactive windows and MLwiN, the back end of R2MLwiN, uses a menu-driven interface for model specifications. On the contrary, bayesmh is relatively difficult for novel users as it uses Stata scripts to define models. Second, most of these software packages offer the default distribution types and parameter values for priors. As exceptions, glmmAK and DPpackage require users to specify the parameter values of

the pre-determined distributions of priors so that the users can specify informative priors based on empirical information. Third, most packages use default Bayesian algorithms that usually are optimized for certain types of models. They also have very good documentations and are well-maintained through frequent updates.

Category C: Software Packages for a Particular Type of Bayesian Multilevel Models

The software packages in category C are tailored for a particular type of Bayesian multilevel models. Specifically, Bayesthresh (Correa & de Sousa Bueno Filho, 2015) is developed to deal with ordinal categorical responses; bayesSurv (Komárek, 2017) is capable of analyzing survival regression models with flexible error and random effects distributions; mlirt (Fox, 2007) is for multilevel unidimensional item response theory (IRT) modeling for dichotomous or polytomous data; mirt (Chalmers, 2012) can estimate either unidimensional or multidimensional latent trait models with random effects under the IRT paradigm for either dichotomous or polytomous responses; bspmma (Burr, 2012) is designed for meta-analysis; ctsem (Driver, Voelkle, & Oud, 2017) is for multivariate continuous (and discrete) time dynamic modelling; and SpatialExtremes (Ribatet, 2017) and spatial.gev.bma (Lenkoski, 2014) are both for modeling spatial data.

The software packages in category C do not have much flexibility in terms of prior types and options of Bayesian algorithms. They use lme4 or lme4-like formulas or routine functions with pre-determined prior types and default values for model specification. Most of the software packages have very good documentations, but not as good as packages in the other two categories in terms of updates and maintenance. For example, mlirt currently only has available releases for R version ≤ 2.15 and bspmma did not update after 2012.

Recommendations and Suggestions

Recommendations to Users

There is no universally best software package for Bayesian multilevel modeling. Instead, a user should choose a program based on their research purpose and goal. For example, if a user is interested in developing new multilevel models, learning how to use a general purpose program, such as BUGS or Stan in Categorical A might be beneficial in the long term. This is because this type of programs allows the specification of new and innovative models beyond the existing ones. On the other hand, if the purpose is to conduct a particular kind of multilevel analysis, a program in Category B or C might be a better choice. The syntax of such a program is usually simple and straightforward. Therefore, one can avoid many common mistakes and save time from learning extra technical details and writing complex syntax. In addition, many of the programs are R packages. If one is already familiar with R, an R package for multilevel modeling is a natural choice. Between programs in Category B and Category C, it might be worth the effort to learn one in Category B. Programs in Category B are in general more flexible to specify a prior distribution. Furthermore, it supports more types of multilevel models that can be helpful to a user in future studies.

As shown in Table 1, software packages in Category A are still the most popular choices in practice, regardless of their disadvantage in user-friendliness and their high demanding on understanding the Bayesian methods and programing. One possible reason for this is that packages in the other two categories are still less well known. Packages such as brms, MCMCglmm, R2BayesX, and R2MLwiN were designed specifically for Bayesian multilevel modeling and have friendly user interfaces. We hope this review can help users choose the right packages for their research.

Suggestions to Developers

As researchers continue to propose new Bayesian algorithms such as INLA within MCMC for spatial models (Gómez-Rubio & Palmí-Perales, 2017), new software packages or programs are going to be developed. From a user perspective, incorporating the newly developed methods into existing packages can be beneficial. However, it would require additional efforts for developers.

Particularly for multilevel modeling, lessons can be learned from software development for structural equation modeling (SEM). For example, many different SEM programs are available, such as Mplus (Muthén & Muthén, 1998-2017), EQS (Byrne, 1994), LISREL (Jöreskog & Sörbom, 1996), SAS CALIS (SAS Institute Inc., 2017), R package lavaan (Rosseel, 2012), and WebSEM (Zhang, Yuan, & Mai, 2012-2017). Although each program is developed independently, they use the same path diagram notations and mostly the reticular action modeling (RAM) notations (McArdle, 1979). Therefore, it is easy for a user to switch from one program to another. Such practice can be followed in the development of software packages for multilevel modeling.

References

Aguero-Valverde, J. (2013). Full Bayes Poisson gamma, Poisson lognormal, and zero inflated random effects models: Comparing the precision of crash frequency estimates. *Accident Analysis and Prevention, 50*, 289–297.

Asparouhov, T., & Muthén, B. (2010). *Bayesian analysis using Mplus: Technical implementation* (software No. Version 3). www.statmodel.com. Retrieved from https://www.statmodel.com/download/Bayes3.pdf

Austin, P. C. (2010). Estimating multilevel logistic regression models when the number of clusters is low: A comparison of different statistical software procedures. *The International Journal of Biostatistics, 6* (1), Article 16.

Austin, P. C. (2017). A tutorial on multilevel survival analysis: methods, models and applications. *International Statistical Review*, *85* (2), 185–203.

Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2014). *Hierarchical modeling and analysis for spatial data*. www.crcpress.com: CRC Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823* .

Bolker, B., Skaug, H., Magnusson, A., & Nielsen, A. (2016). glmmADMB: Generalized linear mixed models using AD Model Builder (version 0.8.3.3) [Computer software manual]. Retrieved from https://rdrr.io/rforge/glmmADMB/

Brezger, A., & Lang, S. (2006). Generalized structured additive regression based on Bayesian p-splines. *Computational Statistics & Data Analysis, 50* (4), 967–991.

Broström, G., & Holmberg, H. (2011). Generalized linear models with clustered data: Fixed and random effects models. *Computational Statistics & Data Analysis*, *55* (12), 3123–3134.

Brown, P. E., & Zhou, L. (2010). MCMC for generalized linear mixed models with glmmBUGS. *The R Journal*, *2* (1), 13–17.

Browne, W. J. (2017). MCMC estimation in MLwiN (version 3.01) [Computer software manual]. Retrieved from https://www.bristol.ac.uk/cmm/media/migrated/2-31/mcmc-print.pdf

Buerkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80* (1), 1–28.

Burr, D. (2012). bspmma: An R package for Bayesian semiparametric models for meta-analysis. *Journal of Statistical Software*, *50* (4), 1–23.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming*. Sage.

Carroll, R., Lawson, A., Faes, C., Kirby, R., Aregay, M., & Watjou, K. (2015). Comparing INLA and OpenBUGS for hierarchical Poisson modeling in disease mapping. *Spatial and Spatio-temporal Epidemiology*, *14*, 45–54.

Chagneau, P., Mortier, F., Picard, N., & Bacro, J.-N. (2011). A hierarchical Bayesian model for spatial prediction of multivariate non-Gaussian random fields. *Biometrics*, *67* (1), 97–105.

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48* (6), 1–29.

Chen, F., Brown, G., & Stokes, M. (2016). *Fitting your favorite mixed models with proc MCMC* (software No. SAS5601-2016). SAS Institute Inc. Retrieved from https://support.sas.com/resources/papers/proceedings16/SAS5601-2016.pdf

Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The*

American Statistician, *49* (4), 327–335.

Correa, F. M., & de Sousa Bueno Filho, J. S. (2015). Bayesthresh: A package for categorical data analysis using Bayesian inference (R package version 2.0.1) [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=Bayesthresh

Correia, M. G., Tura, B. R., & Lanzillotti, R. S. (2016). Developing graphical user interface for Bayesian statistics applied to mixed treatment comparison on R commander. *Value in Health*, *3* (19), A91.

Dedrick, R. F., Ferron, J. M., Hess, M. R., Hogarty, K. Y., Kromrey, J. D., Lang, T. R., . . . Lee, R. S. (2009). Multilevel modeling: A review of methodological issues and applications. *Review of Educational Research*, *79* (1), 69–102.

Dorie, V. (2015). blme: Bayesian linear mixed-effects models (R package version 1.0-4) [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=blme

Driver, C., Voelkle, M., & Oud, H. (2017). ctsem: Continuous time structural equation modelling (R package version 2.5.0) [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=ctsem

Fox, J.-P. (2007). Multilevel IRT modeling in practice with the package mlirt. *Journal of Statistical Software*, *20* (5), 1–16.

Gelfand, A. E. (2000). Gibbs sampling. *Journal of the American Statistical Association*, *95* (452), 1300–1304.

Gelman, A., Su, Y.-S., Yajima, M., Hill, J., Pittau, M. G., Kerman, J., . . . Dorie, V. (2016). arm: Data analysis using regression and multilevel/hierarchical models (R package version 1.9-3) [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=arm

Gilks, W. R., Thomas, A., & Spiegelhalter, D. J. (1994). A language and program for complex Bayesian modelling. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *43* (1), 169–177.

Goldstein, H. (2011). *Multilevel statistical models* (Vol. 922). United Kingdom: John Wiley & Sons.

Goldstein, H., Healy, M. J., & Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine*, *13* (16), 1643–1655.

Gómez-Rubio, V., & Palmí-Perales, F. (2017). Spatial models with the integrated nested Laplace approximation within Markov Chain Monte Carlo. *arXiv preprint arXiv:1702.03891*, 1-26.

Gómez-Rubio, V., & Rue, H. (2017). Markov Chain Monte Carlo with the integrated nested Laplace approximation. *arXiv preprint arXiv:1701.07844* , 1–19.

Goudie, R. J., Turner, R. M., De Angelis, D., & Thomas, A. (2017). Massively parallel MCMC for Bayesian hierarchical models. *arXiv preprint arXiv:1704.03216* , 1–25.

Grant, R. L., Carpenter, B., Furr, D. C., Gelman, A., et al. (2017). Introducing the StataStan interface for fast, complex Bayesian modeling using Stan. *Stata Journal*, *17* (2), 330–342.

Grant, R. L., Furr, D. C., Carpenter, B., & Gelman, A. (2016). Fitting Bayesian item response models in Stata and Stan. *arXiv preprint arXiv:1601.03443* , 1–12.

Guo, J., Gabry, J., Goodrich, B., Lee, D., Sakrejda, K., of Columbia University, T., . . . Niebler, E. (2017). rstan: R interface to Stan (version 2.16.2) [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=rstan

Hadfield, J. D., et al. (2010). MCMC methods for multi-response generalized linear mixed

models: the MCMCglmm R package. *Journal of Statistical Software*, *33* (2), 1–22.

Hall, B. (2016). LaplacesDemon: Complete environment for Bayesian inference (R package version 16.0.1) [Computer software manual]. Farmington, CT. Retrieved from http://www.statisticat.com/laplacesdemon.html

Han, C., & Chaloner, K. (2004). Bayesian experimental design for nonlinear mixed-effects models with application to HIV dynamics. *Biometrics*, *60* (1), 25–33.

Harbord, R. M., Whiting, P., et al. (2009). metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata Journal*, *9* (2), 211-229.

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15* (1), 1593–1623.

Jara, A., Hanson, T., Quintana, F., Mueller, P., & Rosner, G. (2017). DPpackage: Bayesian nonparametric modeling in R (version 1.1-7.2) [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=DPpackage

Jöreskog, K. G., & Sörbom, D. (1996). *Lisrel 8: User's reference guide* (2nd ed.). Lincolnwood, IL: Scientific Software International.

Karabatsos, G. (2017). A menu-driven software package of Bayesian nonparametric (and parametric) mixed models for regression analysis and density estimation. *Behavior Research Methods*, *49* (1), 335–362.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90* (430), 773–795.

Komárek, A., & Lesaffre, E. (2008). Generalized linear mixed model with a penalized Gaussian mixture as a random effects distribution. *Computational Statistics and Data Analysis*, *52* (7), 3441–3458.

Komárek, A. (2017). bayesSurv: Bayesian survival regression with flexible error and random effects distributions (R package version 3.1) [Computer software manual]. Retrieved from http://msekce.karlin.mff.cuni.cz/ komarek

Kreft, I. G., Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Sage.

Leckie, G., Charlton, C., et al. (2013). Runmlwin: A program to run the MLwiN multilevel modelling software from within Stata. *Journal of Statistical Software, 52* (11), 1–40.

Lee, K. J., & Thompson, S. G. (2008). Flexible parametric models for random-effects distributions. *Statistics in Medicine, 27* (3), 418–434.

Lenkoski, A. (2014). spatial.gev.bma: Hierarchical spatial generalized extreme value (GEV) modeling with Bayesian model averaging (BMA) (R package version 1.0) [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=spatial.gev.bma

Lindgren, F., & Rue, H. (2015). Bayesian spatial modelling with R-INLA. *Journal of Statistical Software, 63* (19), 1–25.

Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: Evolution, critique and future directions. *Statistics in Medicine, 28* (25), 3049–3067.

Luo, Y., & Jiao, H. (2017). Using the Stan program for Bayesian item response theory. *Educational and Psychological Measurement, 77*, 1–25.

Martin, A. D., Quinn, K. M., & Park, J. H. (2011). MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software, 42* (9), 1–21.

McArdle, J. (1979). Reticular analysis modeling (RAM) theory: The simplicity and generality of structural equations. In *The 87th Annual Meeting of the American Psychological Association.* New York, NY, USA.

Muthén, L., & Muthén, B. (1998-2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén

& Muthén.

Neal, R. M. (1993). *Probabilistic inference using Markov chain Monte Carlo methods* (Tech. Rep. No. CRG-TR-93-1). Toronto, Ontario, Canada: Department of Computer Science, University of Toronto. Retrieved from http://bayes.wustl.edu/Manual/RadfordNeal.review.pdf

Neal, R. M. (2003). Slice sampling. *The Annals of Statistics*, *31* (3), 705–741.

Pang, X. (2010). Modeling heterogeneity and serial correlation in binary time-series cross-sectional data: A Bayesian multilevel model with AR (p) errors. *Political Analysis*, *18* (4), 470–498.

Peugh, J. L., & Enders, C. K. (2005). Using the SPSS mixed procedure to fit cross-sectional and longitudinal multilevel models. *Educational and Psychological Measurement*, *65* (5), 717–741.

Plummer, M. (2017). JAGS version 4.3.0 user manual [Computer software manual]. Retrieved from https://martynplummer.wordpress.com/2017/07/18/jags-4-3-0-is-released/

Plummer, M., Stukalov, A., & Denwood, M. (2016). rjags: Bayesian graphical models using MCMC (R package version 4-6) [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=rjags

Ribatet, M. (2017). SpatialExtremes: Modelling spatial extremes (R package version 2.0-5) [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=SpatialExtremes

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48* (2), 1–36.

Rossi, P. (2017). bayesm: Bayesian inference for marketing/micro-econometrics (R package

version 3.1-0.1) [Computer software manual]. Retrieved from http://www.perossi.org/home/bsm-1

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *71* (2), 319–392.

Rue, H., Martino, S., Lindgren, F., Simpson, D., & Riebler, A. (2013). R-INLA: Approximate Bayesian inference using integrated nested Laplace approximations [Computer software manual]. Retrieved from http://www.r-inla.org

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., & Lindgren, F. K. (2017). Bayesian computing with INLA: A review. *Annual Review of Statistics and Its Application*, *4*, 395–421.

SAS Institute Inc. (2017). The CALIS procedure. In *SAS/STAT® 14.3 User's Guide*. Cary, NC: SAS Institute Inc.

Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. (2003). WinBUGS user manual (version 1.4) [Computer software manual]. Retrieved from https://www.mrc-bsu.cam.ac.uk/wp-content/uploads/manual14.pdf

Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. (2014). OpenBUGS user manual (version 3.2.3) [Computer software manual]. Retrieved from http://www.openbugs.net/Manuals/Manual.html

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64* (4), 583–639.

Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: A package for running WinBUGS

from R. *Journal of Statistical Software*, *12* (3), 1–16.

Umlauf, N., Adler, D., Kneib, T., Lang, S., & Zeileis, A. (2015). Structured additive regression models: An R interface to BayesX. *Journal of Statistical Software*, *63* (21), 1–46.

Wang, K. (2016). Linear and non-linear mixed models in longitudinal studies and complex survey data. *Journal of Biometrics and Biostatistic*, *7* (2), 290–291.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, *11* (Dec), 3571–3594.

West, B. T., & Galecki, A. T. (2011). An overview of current software procedures for fitting linear mixed models. *The American Statistician*, *65* (4), 274–282.

Yarkoni, T., & Westfall, J. (2016). *Bambi: A simple interface for fitting Bayesian mixed effects models* (software No. 10.17605/OSF.IO/RV7SN). Open Science Framework. Retrieved from https://osf.io/rv7sn/

Zhang, Z. (2014). Webbugs: Conducting Bayesian statistical analysis online. *Journal of Statistical Software*, *61* (1), 1–30.

Zhang, Z. (2016). Modeling error distributions of growth curve models through Bayesian methods. *Behavior Research Methods*, *48* (2), 427–444.

Zhang, Z., Hamagami, F., Lijuan Wang, L., Nesselroade, J. R., & Grimm, K. J. (2007). Bayesian analysis of longitudinal data using growth curve models. *International Journal of Behavioral Development*, *31* (4), 374–383.

Zhang, Z., Parker, R., Charlton, C. M., & Browne, W. J. (2016). R2MLwiN: A package to run MLwiN from within R. *Journal of Statistical Software*, *72* (10), 1–43.

Zhang, Z., Yuan, K.-H., & Mai, Y. (2012-2017). WebSEM: Structural equation modeling online

[Computer software manual]. Retrieved from https://websem.psychstat.org

Table 1

Utilization of Software Packages for Bayesian Multilevel Modeling between 2012 and 2016

Package	Counts	Proportion	Fields
WinBUGS/R2WinBUGS	39	42%	Multidiscipline
R/S-plus programs	10	11%	Multidiscipline
OpenBUGS	9	10%	Multidiscipline
Matlab programs	7	8%	Neuroscience
C/C++/Python programs	6	7%	Disease Preventing/Medical research/Clinical trials/Geography
R–INLA (INLA)	5	5%	Epidemiology/Biology
JAGS/R2jags	5	5%	Macular Degeneration/Seismology
SAS MCMC	4	4%	Ecology/Cancer research/Clinical Psychology
Stan/rstan	2	2%	Biometrics
ADMB/ADMB-RE	2	2%	Fisheries science/Ecology
Stan brms	1	1%	Psychology
MLwiN	1	1%	Social Mobility
spatial.gev.bma	1	1%	Environmetrics

Note. The data are based on a collection of 92 applied studies published between 2012 and 2016.

Table 2

Technical Options and Features for Modeling

Package	Prior Types	Core Bayesian Algorithms	Parallel Computing	Running Environment	Bayesian Engine	Programming Languages
Category A						
R2WinBUGS/R2OpenBUGS	Flexible	Gibbs, MH, Slice	No	R	BUGS	Component Pascal, R
rjags	Flexible	(Adaptative) Gibbs, MH, Slice	No	R	JAGS	C++, R
MultiBUGS	Flexible	Gibbs, MH, Slice	Single chain parallel: multiple cores/APIs	Stand alone	OpenBUGS	Component Pascal
SAS MCMC	Flexible	MH, IS	Multiple threads	SAS	SAS MCMC	C
R-INLA (INLA)	Flexible	INLA	Multiple threads	R	C++ GMRFLib	C++, R
LaplacesDemon	Flexible	HARM, LaplaceA, INCA, HMC	Multiple cores	R	LaplacesDemon	R
rstan	Flexible	HMC, NUTS	Multiple cores	R	Stan	C++, R
Categroy B						
brms	Flexible	HMC, NUTS	Multiple cores	R	Stan	C++, R
MCMCglmm	Fixed	Gibbs, MH, RWM, Slice	—	R	MCMCglmm	C/C++, R
glmmBUGS	Fixed	Gibbs, MH, Slice	No	R	R2WinBUGS	C++, R
glmmAK	Limited	ARG	—	R	glmmAK	R
blme	Limited	—	—	R	blme	R
Bambi	Flexible	NUTS, HMC	Multiple cores / GPUs	Python	Python PyMC3	Python
R2BayesX	Fixed	Gibbs, RW, P-splines	R parallel	R	BayesX	C++, R
MCMCpack	Fixed	a generic RWM	—	R	R MCMC	C++, R
DPpackage	Fixed	MH, IWLS: Model-specific	—	R	DPpackage	Fortran
BayesReg	Fixed	Gibbs, ARWMH	—	Matlab	Matlab BayesReg	Matlab
bayesmh	Limited	Gibbs, MH	No	Stata	Stata MCMC	C
bayesm	Fixed	Gibbs, MH, RWM	—	R	bayesm	C++, R
R2MLwiN	Flexible	Gibbs, MH, SMVN	GPUs	R	MLwiN MCMC	MLwiN macro, R
glmmADMB	Fixed	RWM	No	R	AD Model Buider	C++, R
Mplus Bayes	Limited	Gibbs, MH	Multiple cores	Stand alone	Mplus Bayes	Java
arm	Fixed	—	—	R	lme4::mcmcsamp	R
Categroy C						
Bayesthresh	Fixed	—	—	R	Bayesthresh	R
bayesSurv	Fixed	Slice, ARG	—	R	BayesSurv	R
mlirt	Fixed	Gibbs, MH within Gibbs	—	R	Fortran IMSL	Fortran, R
mirt	Fixed	MHRM	R parallel	R	mirt	C++, R
bspmma	Fixed	—	—	R	bspmma	R
ctsem	Fixed	NUTS, HMC	Multiple cores	R	Stan	C++, R
SpatialExtremes	Fixed	Gibbs	—	R	SpatialExtremes	R
spatial.gev.bma	Fixed	MH	—	R	SpatialExtremes	R

Note. "—" = Not sure. MH = Metropolis–Hastings. IS = Importance sampling. INLA = Integrated nested Laplace approximation. LaplaceA = Laplace approximation. HARM = Hit-and-run Metropolis. INCA = Interchain adaptive. HMC = Hamiltonian Monte Carlo. NUTS = No-U-Turn sampler. RWM = Random walk Metropolis. ARWMH = Adaptive random-walk Metropolis–Hastings. SMVN = Structured multivariate normal. ARG = Adaptive rejection Gibbs. MHRM = Metropolis-Hastings Robbins-Monro. API = Application programming interface. GPU = Graphics processing unit.

Table 3

Model Specification, Algorithms Selection, Documentation, Installation, Open source, and Updates

Package	Model Specification		Algorithm Default/Custom	Documentation	Installation Dependency/Compiler	Open Source	Published/ Updated
	Tools	Default Priors Type/Value					
Categroy A							
R2WinBUGS/R2OpenBUGS	BUGS script	No/No	No/Yes	Outstanding	WinBUGS/OpenBUGS	GNU, GPL-2	2005/2017
rjags	BUGS script	No/No	No/Yes	Outstanding	JAGS	GNU, GPL-2	2008/2016
MultiBUGS	BUGS script / DAG graph	No/No	No/Yes	Outstanding	No	GNU	2017/2017
SAS MCMC	SAS script	No/No	Yes/Yes	Outstanding	No	No	2008/2017
R-INLA (INLA)	formula, matrix, and routine functions	Yes/Yes	Yes/Yes	Outstanding	C++ compiler	GNU	2011/2017
LaplacesDemon	R script	No/No	Yes/Yes	Very good	No	MIT	2011/2016
rstan	Stan script	No/No	Yes/Yes	Outstanding	StanHeader/C++ compiler	New BSD, GPL	2012/2017
Categroy B							
brms	lime4-like formula, function arguments	Yes/Yes	Yes/Yes	Very good	StanHeader/C++ compiler	GPL >= 3	2015/2017
MCMCglmm	function arguments	Yes/Yes	Yes/No	Very good	C++ compiler	GPL >= 2	2009/2016
glmmBUGS	formula, function arguments	Yes/Yes	Yes/No	Very good	WinBUGS	GPL	2008/2016
glmmAK	routine functions and arguments	Yes/No	Yes/No	Very good	No	GPL-2	2007/2015
blme	lme4 formula, function arguments	Yes/Yes	Yes/No	Very good	No	GPL >= 2	2011/2015
Bambi	lme4 formula, function arguments	Yes/Yes	Yes/No	Good	Python Interpreter	MIT	2016/2017
R2BayesX	formula, function arguments	Yes/Yes	Yes/No	Very good	BayesXsrc	GPL-2	2005/2017
MCMCpack	lme4 formula, function arguments	Yes/Yes	Yes/No	Very good	C++ compiler	GPL-3	2003/2017
DPpackage	lme4-like formula & function arguments	Yes/No	Yes/No	Very good	Fortran compiler	GPL >= 2	2006/2017
BayesReg	template math equations	Yes/Yes	Yes/No	Very good	Matlab compiler	—	2015/2017
bayesmh	Stata script	Yes/Yes	Yes/Yes	Good	Stata	No	2015/2017
bayesm	routine functions and arguments	Yes/Yes	Yes/No	Very good	C++ compiler	GPL >= 2	2006/2017
R2MLwiN	lme4-like formula, function arguments	Yes/Yes	Yes/Yes	Outstanding	MLwiN	GPL >= 2	2012/2017
glmmADMB	lme4-like formula, function arguments	Yes/Yes	Yes/No	Good	AD Model Builder	BSD_2_CLAUSE	2005/2017
Mplus Bayes	Mplus syntax	Yes/Yes	Yes/Yes	Outstanding	No	No	2010/2017
arm	lme4 formula	Yes/Yes	Yes/No	Good	No	GPL >= 3	2007/2016
Categroy C							
Bayesthresh	lme4 formula and function arguments	Yes/Yes	Yes/No	Very good	No	GPL-3	2012/2013
bayesSurv	lme4-like formula, function arguments	Yes/Yes	Yes/No	Very good	No	GPL-3	2009/2017
mlirt	a routine function and arguments	Yes/Yes	Yes/No	Very Good	No	no commercial use	2007/2010
mirt	matrix, routine functions and arguments	Yes/Yes	Yes/No	Very good	C++ compiler	GPL > = 3	2010/2017
bspmma	function arguments	Yes/Yes	Yes/No	Very Good	No	GPL-2	2012/2012
ctsem	function arguments	Yes/Yes	Yes/No	Very good	rstan/C++ compiler	GPL-3	2016/2017
SpatialExtremes	formula, function arguments	Yes/Yes	Yes/No	Very Good	No	GPL >= 2	2009/2017
spatial.gev.bma	function arguments	Yes/Yes	Yes/No	Good	No	GPL	2014/2014

Note. "—" = Not sure. GPL = General Public License. BSD = Berkeley Software Distribution. MIT = Massachusetts Institute of Technology.

Table 4

Supported Multilevel Models by Packages in Category B

Package	Responses							Models					
	Con.	**Bin.**	**Cou.**	**Ord.**	**Mul.**	**Multivariate**	**Zero-inflated**	**Weighted**	**Cross-classified**	**Survival [a]**	**Spatial**	**Additive**	**Customization**
brms	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
MCMCglmm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	No
glmmBUGS	Yes	Yes	No	No	No	Yes	No	No	No	No	Yes	No	Yes
glmmAK	No	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No
blme	Yes	Yes	Yes	No	No	No	No	Yes	Yes	No	No	No	No
Bambi	Yes	Yes	Yes	No	No	No	No	No	Yes	No	No	No	Yes
R2BayesX	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	No
MCMCpack	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	Yes
DPpackage	Yes	Yes	Yes	No	No	Yes	No	No	—	Yes	No	No	No
BayesReg	Yes	Yes	—	Yes	No	—	Yes	Yes	—	Yes	Yes	Yes	No
bayesmh	Yes	Yes	—	—	—	—	—	—	Yes	—	—	—	Yes
bayesm	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	—	—	—	—	No
R2MLwiN	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	—	Yes	No	Yes
glmmADMB	Yes	Yes	Yes	No	No	—	Yes	—	No	—	No	No	No
Mplus Bayes	Yes	Yes	Yes	Yes	—	Yes	—	Yes	Yes	—	No	No	No
arm	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	No	No	No	No

Note. "—" = Not sure. Con. = Continuous. Bin. = Binary. Cou. = Count. Ord. = Ordinal. Mul. = Multinomial. a. By restructuring and transforming the data, standard software packages for multilevel modeling can fit two types of survival models: piecewise exponential (PWE) models and discrete time models (Austin, 2017).